# الحاسوب والصناعة المعجماتية

د. عبد الغني أبو العزم (*)

أصبح استخدام الحاسوب في مجال الصناعة المعجماتية ضرورة علمية لا محيد عنها لأي مشتغل بتأليف المعاجم، وقد فرض نفسه في مجال البحث اللساني والمعجمي والمعجماتي، محدثا بذلك تحولا جذريا في مفهوم المعجم وصناعته وأدوات تحضيره وإنجازه.

ما هي الخصائص الآلية التي أصبحت تتوفر عليها الأداة الحاسوبية؟

يمكن أن نجمل هذه الخصائص في ميزتين أساسيتين:

I- تخزين النصوص وبرمجتها.

II- ترتيب المفردات ورسم الألفاظ وجذورها، لتصبح معالجتها معالجة آلية بعد إدخالها في أنظمة دقيقة مما يسمح بالتطبيق العلمي المباشر لاستعمالاتها المتعددة.

تعد هاتان الميزتان الأساسيتان من التطبيقات العلمية للأداة الحاسوبية في مجال الصناعة المعجماتية، وترتبطان ارتباطاً عضويا بتنظيم بنوك المعطيات، وما تقدمه من معلومات ضخمة، لم يعد بالإمكان الاستغناء عنها في الإنجاز المعجمي والمعجماتي. وقد أصبح من العبث أن يستخدم المعجماتي الأساليب العتيقة في إنجاز معاجم لغوية أو معاجم متخصصة في زمن تحكم سيره الآلات الحاسوبية كأداة لها فعالية حاسمة سرعة وإتقانا.

إن ما أصبح يوفره الحاسوب في مجال المعجماتية يمكن تحديده في النقاط التالية:

- إنجاز معاجم متنوعة ومتخصصة وبدقة متناهية وبأقصى ما يمكن من السرعة.
- توفير الإرهاق والتعب
- ربح مصاريف باهظة.

ولإيضاح القدرة الهائلة والفائقة التي أصبح يتوفر عليها الحاسوب، يمكن الإشارة هنا إلى واجهتين بارزتين:

(*) أستاذ بكلية الآداب امعة ا سن الثا (ع الشق) ا غرب

1. واجهة التخزين

2. واجهة استحضار المعلومات.

**1**– لقد استطاع التطور التكنولوجي في مجال المعلوماتية أن يحقق تراكما في مادة التخزين، إذ أصبح بإمكان الحاسوب أن يخزن ما يزيد على 470 مليارا من الحروف، وعلى هذا الأساس يمكن أن يستوعب مليون كتاب (أي بمعدل مائتي صفحة لكل كتاب).(1)

**2**– لقد ارتبط هذا التخزين بتطور مواز، أي التمكن من التحليل الآلي للنصوص، مما يسمح بإمكان استحضار أكبر عدد ممكن من المعلومات، وهي تقدر بالملايين وفي دقائق معدودة، وأصبح من السهل إيجاد الجملة المطلوبة والمحددة من بين جمل المليون كتاب المخزنة في أقل من عشرين ثانية. وللمزيد من الإيضاح، يمكن أن أضيف في مجال التخزين أن ذاكرات الحاسوب وحسب طاقاتها الاستيعابية قادرة على تسجيل كل كتب دار الكتب المصرية زائد كتب الخزانة العامة بالمغرب. ولا يقف الأمر فقط عند التخزين كما أشرت، إذ أن مؤلفا من 80.000 كلمة إذا كان مسجلا في الذاكرة يمكن أن نحصل على كلماته مباشرة بأي شكل نريده، في تصاريفها أو استعمالاتها المختلفة، منسقة ومرتبة ترتيبا ألفبائيا مع فهرس تواترها، كما يمكن أن نحصل على رقم الصفحة والسطر والفقرة الواردة فيها.

ولقد شرعت بعض المؤسسات العربية وبنوك المعلومات منذ أواخر الثمانينات وبداية التسعينات في الاستفادة من هذه التقنية الحديثة، حيث تم وضع برمجيات دقيقة، من بينها: نظام تصحيح الأخطاء الإملائية والشكلية مع إمكان تغيير مواقع الجمل والكلمات،وإلغاء الفقرات، بالإضافة إلى تغيير أجزاء الكلمات والجمل، وبذلك قد تم التغلب على العديد من الصعاب التي كانت ترهق كواهل المعجماتيين بفضل تقنية الحاسوب.

وما زالت الجهود العلمية تتواصل وذلك من أجل:

- حصر وضبط وتسجيل مختلف المعطيات
- إنجاز برامج إلكترونية.
- طبع ونشر النتائج.

وتبقى المشكلة الأولى أعلاه هي حجر الزاوية، ومن المفيد هنا أن نعرض لبعض المصطلحات والقضايا المعجمية التي أصبحت ترتبط بالإنجاز المعجماتي الآلي مما يساعدنا على فهم آلية اشتغال الحاسوب أو ما يعرف بالمكنز– حسب تسمية الشركة العالمية بالقاهرة– وحدود ما يقدمه من خدمات.

## المعجمية النصية

يعد مصطلح المعجمية النصية من بين المصطلحات الحديثة التي ظهرت مع تطور الحاسوب واتساع قدرته على استيعاب أكبر عدد من النصوص.

ترتكز المعجمية النصية في إنجازاتها على ما يقدمه الحاسوب وتستفيد منه على مستويين:

المستوى الكمي، يتحدد في إطار توفير ومعالجة الأعداد الهائلة من الجذاذات، أي ما يقدمه محلل المداخل من معان وحصرها، وينتج عن هذا الحصر، حصران: حصر المتداول، وحصر المهمل.

يتحول الحاسوب في هذا الصدد إلى وعاء ضخم، سواء في تعامله مع الجرد الموسع للألفاظ، أو في تعامله مع كل الاستعمالات في مجموع النصوص الموثقة بداخله، وهو في هذه الحالة يقدم لمستعمله مادة ضخمة منسقة وشبه منسجمة ومرتبة ترتيبا ألفبائيا، وفي هذه الحالة لم يعد المعجماتي بحاجة إلى أن يمر بكل المراحل التقليدية التي اعتاد على اجتيازها.

يتعامل الحاسوب بطبيعة الحال مع كل النصوص التي زودناه بها، مما يعني أن اختيارها يبقى مرهونا بالأهداف المتوخاة. ومادامت الطاقة الاستيعابية للحاسوب ذات اتساع غير محدود، فمن المفروض أن يضم بالإضافة إلى ما يختزنه من نصوص ومعاجم أعمال المجامع اللغوية، والمؤسسات العلمية، ومكتب تنسيق التعريب والأعمال الفردية المتوفرة. إن إدخال وتخزين كل هذه الأعمال يجعلنا متأكدين أننا نملك على أقل تقدير 95% من مجموع الألفاظ والمصطلحات المتداولة في السوق اللغوية العربية.

يسمح هذا الإنجاز بامتلاك المعرفة والمادة، وللتمكن منهما لا بد من تكاتف جهود اللسانيين والمعجماتيين ورجال المعلوماتية، والمهندسين والمخترعين لمختلف الأجهزة وما يوفرونه من أنظمة وبرامج.

تعتمد المعجمية النصية على مدى طاقة الحاسوب الاستيعابية، وهو ما تم التغلب عليه بفضل التقدم الهائل الذي حقق اتساع ذاكرته والإمكانات التي يوفرها في تعامله مع النصوص، وهو ما يؤسس لأربع مجالات أساسية:

I– أساسيات المعطيات — **Les Bases de données**

II– أساسيات المعطيات النصية — **Les bases de données textuelles**

III– أساسيات المعطيات المعجماتية — **Les bases de données lexicographiques**

IV– أساسيات المعطيات القاموساتية — **Les bases de données dictionnairiques**

ماذا يقصد بهذه الأساسيات الأربع؟ وما هي أهميتها في الإنجازات المعجماتية؟ هذا ما سنحاول إيضاحه اعتمادا على العديد من الدراسات التي أنجزت في هذا الصدد(2).

## I – أساسيات المعطيات:

وفرت النتائج الحالية للتطور التكنولوجي في ضوء ما سبقت الإشارة إليه مجالا واسعا ورحبا للدراسات والأبحاث المعجمية والمعجماتية، وبذلك اكتست أساسيات المعطيات أهمية قصوى حيث تهم كل المتخصصين في شتى المجالات وأصبح بإمكانهم استغلال أقصى ما يمكن من الوثائق بفعالية كبيرة. وهنا يجب التمييز بين أساسيات المعطيات، وبنوك المعطيات، مع العلم أن هذا المصطلح الأخير حسب برنادر كيمادا قد صار تداوله لغويا يحمل قيمة خصوصية، ولإزالة الالتباس يعطي تعريفين إجرائيين لمصطلح أساسيات المعطيات، يحدد التعريف الأول في كونه يعني نصا مُحَوْسَباً مكونا بارتباط مع مجموع المعطيات، وقائع أو معلومات لها طبيعة مشتركة مسجلة على الحاسوب حسب قواعد نظامية يمكن الوصول إليها، وتخضع لبرنامج أو برامج للتحليل ومهيأة للتساؤل ومكيفة بالخصوص في ضوء تلك القواعد.

بينما بنك المعطيات له بعد إضافي، ويناسب مؤسسة وثائقية معلوماتية( Institution documenataire informatisée ) تدير وتستغل نصا أو عدة نصوص من نموذج الأساس لصالح جماعة المستعملين، و له أيضا أهداف نوعية منها جمع تشكيلات من المعلومات التراكمية، وبما يمكن من الشمولية والاستقصاء تستجيب لما يطلب منه، وهو بهذا الشكل عليه أن يساهم في حل مشاكل التسيير الإداري والقانوني (3).

وفي هذه الحالة وكما يلاحظ أيضاً برنارد كيمادا تبدو فكرة قاعدة المعطيات تافهة في مبدئها إلا أنها بالنسبة للمعجماتي تشكل وظيفة أساسية في الإنجازات المعجمية والمعجماتية المنتظرة لما توفره من مواد مزدوجة الأهداف، فمن جهة هناك المدونة (نصوص أو عناصر نصية معيارية أو نصوص طبيعية أو موضوعات مداخل المعجم) حيث يرتبط تسجيلها أو تهييئها بالاستعمالات اللاحقة، ومن جهة أخرى فإن مجموعة الخورزمات الموضوعة للتساؤل (برامج، معاجم آلية، أنظمة وثائقية) قد تصبح ممكنة انطلاقا من الاستشارات والتحاليل التلقائية للمدونات، مع العلم أن درجتها المعقدة هي التي تحدد باستمرار قيمتها، ومستوى التعامل مع كل النتائج المحصل عليها.

وينبغي أن نشير هنا إلى أن هذه التحاليل في الأعمال المعجماتية يمكن أن تكون ذات اتجاهين:

1. النمط التعاقبي، ويعد هذا النمط من الأنماط الأكثر بساطة، إذ يسمح بالكشف عن المتواليات أو عن صلة ترابط المفردات حسب نظام الخطاب (توزيعات وتوافقات ثنائية).

2. النمط العلائقي: وهو ما يسمح بإيجاد العناصر المنتمية إلى أقسام أو خانات أو إلى جذور معجمية.

لقد انتقلت هذه الإجراءات العملية التي كان يقوم بها المعجماتي يدويا إلى التطبيق الآلي، ويتمثل الفرق بداهة في مستوى الأداء والإتقان والضبط والسرعة، وهذا تطور آخر نوعي يدخل في سياق مجمل مراحل التاريخ الطويل لتطور المعجمية والمعجماتية (4).

وقبل أن نختم الحديث عن أساسيات المعطيات وإجرائياتها، وفي ضوء التمييز الذي أشرنا إليه فيما يخص بنك المعطيات، نود هنا أن نبرز ماهية ووظائف ما يعرف ببنك النصوص، وبنك الكلمات وتحديد أدوارهما على قاعدة أساسيات المعطيات، وهذا التحديد ضروري لتمييز عمل الأجهزة المعلوماتية والوثائقية المختلفة.

**1- بنك النصوص:** تختلف وظيفة بنك النصوص حسب الأهداف الموضوعة، وهو يقوم أساسا حسب تعريف كيمادا بتجميع نصوص تتضمن كل أنواع وأشكال الوثائق المكتوبة أو المنقولة، وتخضع للتعيين والتسجيل حسب معايير مكيفة في ضوء حاجات مستعمليها. وبما أن إشكالية بنك النصوص تتجاوز الكفاءات والاهتمامات النوعية للمعجميين والمعجماتيين فإن طبيعة الوظائف التي يؤديها تمس كل التخصصات، وقد أصبحت متعددة مما يفرض وضع اختيارات علمية ومنهجية وتقنية بمجرد ما يتم تكوين المدونات على مستوى اختيارات النشر والتنسيق والضبط وتنظيم أسس العمل. إن المستفيد الأكبر من بنك النصوص، هو المعجماتي إذا ما تم توسيع نطاق مخزونه. لهذا يشكل بنك النصوص بالنسبة إليه أداة عملية تجعله محيطا بدقة بشتى النصوص التي يود الرجوع إليها عند الطلب والحاجة، وعندما يضع خطة لإنجاز أي عمل معجماتي.

وعلى سبيل المثال وفي ضوء بنك النصوص يمكن إنجاز:

- المعجم العربي الشامل الذي يضم كل مفردات اللغة العربية ومصطلحاتها قديما وحديثا.
- المعجم التاريخي.
- إنجاز معاجم متخصصة في شتى المجالات.
- معاجم جغرافية وتاريخية وفنية.

هذا بالإضافة إلى معاجم المترادفات والأضداد.

وإذا كان بنك النصوص يقدم للمعجماتي خدمات هائلة، ويضع بين يديه مختلف السياقات في شتى العلوم والفنون فإن النصوص التي يخزنها تحتاج إلى تحليل ومعالجة ليتأكد من صحتها ودقة تعابيرها، وهذا ما تقدمه أساسيات المعطيات النصية وفق البرامج المعدة لهذا الغرض، وبالتحديد ما يتعلق بالإنجاز المعجماتي، مع الأخذ بعين الاعتبار كل التطورات التكنولوجية، وما أسفرت عنه من نتائج تقنية أدت إلى تطوير أنظمة وبرامج خاصة ببناء

أساسيات المعطيات النصية. ومن بين هذه البرامج نشير إلى:

القراءة الضوئية،وهي ما يعرف حاليا بالتمييز الضوئي للحروف O ) Optical character Recognition (C R

لقد أحدثت القراءة الضوئية انقلابا جذريا في بناء أساسيات المعطيات النصية، وهي تعد من أهم المنجزات الإلكترونية حيث تم تجاوز عمليات الإدخال اليدوي للمعطيات من خلال لوحات المفاتيح.

إن أهم ما تقدمه القراءة الضوئية هو إمكان إحداث نظام البحث والاسترجاع للنصوص مع قدرتها المتميزة في معالجة اللغات، مما سيفتح مجالا رحباً لمعالجة النصوص بدقة متناهية وبأقل التكاليف.

وبالإضافة إلى ذلك، يمكّن هذا النظام من إدخال عدد لانهائي من محتويات النصوص في وقت وجيز بواسطة المصور الضوئي، أي ما يعرف الآن بـ (Scanner) مما يمكّن أي برنامج تنسيق للتعامل معها ومعالجتها بسهولة تامة كأي نص تم إدخاله من خلال لوحة المفاتيح.

تشتغل القراءة الضوئية في بيئة النوافذ مما يسهل عملها حيث تتمكن من التشغيل بسهولة مع إمكان الاتصال بتطبيقات اللغات الأجنبية.

ونشير هنا إلى أن مؤسسة (العالمية للإلكترونيات) التي وقفت شخصيا على إنجازاتها بالقاهرة استطاعت أن تطور برنامج القراءة الضوئية تطويرا هائلا حيث أصبحت "تتعامل مع الطبيعة الخاصة للغة العربية المتميزة بصعوباتها المتمثلة في انحناءات الحروف واختلاف شكل الحرف،وبالتالي ضرورة تقسيم الكلمة إلى حروف أولا وتمييزها ثم قراءتها، وهما عمليتان في مقابل العملية الأخيرة فقط في اللغات اللاتينية، وقد اقتضى هذا نوعا خاصا من المعالجة لقراءة جميع أنواع الخطوط وأحجامها باستخدام تقنية(Contour Analysis) فأصبح البرنامج قادرا على قراءة جميع أنواع الخطوط العربية المتاحة في الأسواق" (5).

وبالإضافة إلى هذا كله يتضمن البرنامج ملفات عمل آلية لخطوط معينة تفيد المستخدم عندما يرغب في إدخال عدد هائل من الوثائق، أو النصوص، بعدد قليل من الخطوط. وهذا ما يسمح أيضاً بتصحيح الأخطاء، أي إتاحة وضع مجموعة كبيرة من الوثائق، تفيد المستخدم عندما يرغب في إدخال عدد هائل من الوثائق أو النصوص، أو وضعها في المصور الضوئي (Scanner ) وهذا يسمح أيضاً بتصحيح الأخطاء، مع العلم أن النصوص المصورة بالمصور الضوئي يقوم البرنامج بقراءتها مباشرة دون تدخل خارجي.

## 1- تحليل النصوص:

يخضع ما تم تخزينه من نصوص لعمليات تحليلية لجميع المفردات الواردة فيها تحليلا صرفيا كاملا (اشتقاقيا وتصريفيا) حيث يتم الحصول على بدائلها الصرفية، وهذا النظام يعرف بـ (المحلل الصرفي). حيث يتولى تحليل أي مفردة تحليلا صرفيا ويقدم لك جذرها كما يردّها إلى استعمالاتها الصرفية الممكنة بسوابقها ولواحقها مع إبراز موازينها وصيغها الصرفية ومشتقاتها المختلفة، وما يلحقها من تغيير في حالة الإفراد والتثنية والجمع.

وحسب ما وقفت عليه شخصيا في مقر الشركة العالمية للإلكترونيات بالقاهرة أن هناك تقدما هائلا تم تسجيله فيما يخص التحليل الصرفي واستراتيجية التصحيح، ولأعطي مثالين عن ذلك:

1. في حالة إدخال مفردة مجردة عن النص مثل (قال) إلى النظام الآلي للتحليل الصرفي، فإنه يقدم كل بدائلها الصرفية.
2. وفي حالة إدخال مفردة واردة في نص ما مثل " أحدث" إلى النظام الآلي للتحليل الصرفي، فإنه يقدم كل بدائلها الصرفية.

وتكمن أهمية هذا الإنجاز في تمكينه من استيعاب مجمل الكلمات القديمة والمهجورة عندما يتم رصدها وتخزينها ليتمكن من تحليلها تحليلا صرفيا، هذا بالإضافة إلى استيعاب كلمات أجنبية سواء كانت ألفاظاً ومصطلحات معربة، أو أماكن وأعلاما، مثل رأسمالية وماركسية وبنيوية واستراتيجية واستقراطية، وروما و واشنطن و موسكو.كل هذه الكلمات يمكن تصنيفها تحت خانة الكلمات غير قابلة للاشتقاق وهذه الخانة تتفرع إلى خانات منها:

1. الكلمات التي لا جذر لها وتؤخذ منها المشتقات المختلفة كالمبنيات من الحروف والأسماء مثل حروف الجر وأسماء الإشارة وأسماء الحروف والأفعال الجامدة مثل ليس وعسى والألفاظ المنحوتة، مثل البسملة والحمدلة وتضاف مجمل الكلمات التي لا تحلل تحليلا صرفيا مثل لفظ الجلالة (الله).
2. الأعلام العربية، لأن تخزينها يسمح بتصنيفها مما يجعل إمكان البحث أمراً ميسوراً، وتمكن الباحث من الوقوف على معلومات إضافية.
3. الأعلام الأجنبية: وتصنف بدورها في خانة مع معلومات إضافية عنها.
4. الأعلام المجهولة: وهي بدورها يتم تصنيفها سواء كانت متداولة ولا يعرف أصلها اللغوي أو عامية.
5. الأخطاء الشائعة: إن من بين ما يقدمه التخزين القائم على التصنيف أنه يمدك بالمفردة المبحوث عنها في استعمالاتها الصحيحة والسليمة وينبه إليها، مثال على ذلك:

أ- المصادر الخماسية والسداسية- يتم التنبيه إلى أنها لا تكتب بهمزة قطع.

**ب–** إنشاء الله – يتم التنبيه إلى أن (إن) حرف الشرط لا تكتب متصلة.

**ج–** وفي حالـة ورود كـلمة في بيـت شـعري خضـع للضـرورة الشعرية، يتم أيضاً التنبيه إلى ذلك، مثل: القائمينا) (الفضا).

**د–** وفي حالة ورود كلمة عامية تتم الإشارة إلى ذلك.

تخـرج كـل المفـردات المشـار إليهـا آنفا عن نطاق النظام الآلي للتحليل، لأنها لا تخضع للقواعد القياسية المزودة بها، لهذا وضعت في إطار ملف يقدمها البرنامج موافقة لأغراض كاتبيها.

**هــ–** وفي حالـة مـا إذا تم عـرض مصـطلحات مركبة من كلمتين مثل (الاجتماعي– الثقافي) أو(الحوار الأوروبي– العربي) على النظام الآلي للتحليل الصرفي، فإنه لا يقبلها لكونه لا يحلل كلمتين معا في آن واحد. ولمعالجة هذا الإشكال تم تخصـيص مـلف لهـذا الـنوع مـن المصطلحات بحيث يمكن للباحث أو المعجماتي أن يقف بسهولة تامة على تركيبها ودلالتها في ملفها وحسب برمجتها.

وهـناك جـانب أساسـي يتميز به النظام الآلي للتحليل الصرفي يتجلى في قدرته على التصحيح، فهو يتقبل مجمل النصـوص عـند إجراء عملية التخزين،إلا أنه بإمكانه أن يميز كل الكلمات الخاطئة سواء أكانت أخطاء إملائية أم لغوية أم طـباعية، وسـواء تعـلق الأمر بمفردة بها خطأ أو أكثر مثل (إقتصاد)(إجل)(كتابت) أو كلمة متشابكة عندما يتم إسقاط المسـافات الـتي ينـبغي أن تكـون بيـنها مثل (منذز من بعيد) (الر باط) نيو يورك) أو كلمة منفصلة، أي انفصال أجزائها أثـناء إدخـال النصـوص بسـبب إقامـة المسـافات، إذ مـن المفـترض فيهـا الاتصال (علا مات مصا در فيد يو كهرو مغناطيسي).

ويمكـن القـول في هـذا الصـدد إن هـناك اسـتراتيجية دقيقـة وضـعت بدقة متناهية لمعالجة هذا النوع من الأخطاء وجعل النصوص المخزنة تخلو من الأخطاء مهما كان مصدرها إملائية أو لغوية أو طباعية.

## 2– بنك الكلمات :

يوجـد إلى جـانب بـنوك النصـوص حـلقة رئيسـية بـل محورية تتجلى في بنوك الكلمات، إذ أنها تكتسي أرضية خصـبة لأي مشـروع معجماتي، وأصبحت هذه البنوك متداولة بين أكبر عدد ممكن من الناس وتضع بين يدي السائل ما يطلـبه وفي أقصـر وقت ممكن، مع العلم أن مهمة بنوك الكلمات لا تنحصر في مجرد التخزين أو الاستشارة فحسب، بل لهـا وظيفة تقنية واضحة تتحدد في استخدام أو معالجة وتدبير معجم أو معاجم ممعلمة. وقد تأسست هذه البنوك في عدد كـبير مـن البـلدان وأغلـبها كـان يهـتم بموضـوع المصطلحات منذ أواخر الستينات وبداية السبعينات، إلا أنها مع تطور

الطاقة الاستيعابية للحاسوب وسعت نطاق اهتماماتها لتشمل أيضاً المفردات اللغوية، ومن بين هذه البنوك (6):

- بنك المصطلحات في مجال الترجمة أوتاوا.
- (تيرميوم) Termium
- نورماتيرم بباريس Normaterm
- بنك المصطلحات التابع لشركة سيمنز.
- بنك المعلومات المنمطة بألمانيا.
- يوروديكاتوم بلوكسمبورغ Eurodicautom.
- باسم بالمملكة العربية السعودية.
- مكنز الشركة العالمية للإلكترونيات بالقاهرة.
- المعربي Lexar بمعهد الدراسات والأبحاث للتعريب بالرباط .

وهي تؤدي وظيفتها بشكل عادي ولا تنحصر فقط في إطار العمل المعجمي والمعجماتي، بل تمس كافة التخصصات حيث تزود المؤسسات العلمية والمعاهد ومراكز الأبحاث والشركات بما يحتاجه من كلمات ومصطلحات في كل المجالات، وعلى سبيل المثال، إذا ما طلب السائل مفردات متعلقة بالزراعة فإن البنك يمده بها.

وفيما يخص العمل المعجماتي فهو جزء مستقل في إطار ارتباطه العام ببنوك الكلمات، لهذا ينبغي أن يخضع لمعايير محددة انطلاقا منها، مما يدعو إلى:

- إيجاد الرصيد اللغوي الأساسي للغة العربية الذي يخضع إنجازه للبحث.
- إيجاد الرصيد اللغوي لمعجم الطفل، المعجم النواة.
- إيجاد أرصدة لغوية مرحلية أي لكل سنة من سنوات التعليم الأساسي.
- إيجاد الرصيد اللغوي الوظيفي للتعليم الأساسي.

إن هذه الأرصدة بهذا المفهوم وفي اعتمادها على بنوك الكلمات هو ما يجنّب السقوط في الخطإ، ونحن نضع مثلا مشروع المعجم العربي الحديث. إلى جانب هذه الأهمية فإن الأرصدة اللغوية المنجزة في ضوء مجمل النصوص المدرسية هي مادة خام لمؤلفي الكتب المدرسية والمعاهد التربوية.

وبهـذه الصـورة سيتمكن الحاسوب أو ما يعرف بالمكنـز من تأدية خدمة هائلة وسيصبح مرجعاً أساسيا ليستجيب لأي تساؤل. وهذه المهمة مطروحة على الباحثين واستقصاء أعمال المجامع اللغوية والمعاهد العلمية التي تشتغل في مجال المصطلحات والمفردات الحديثة، ولكي يصبح بنك الكلمات مزودا بأقصى ما يمكن من الكلمات.

بعدمـا أوضـحنا الإمكانـات الـتي أصبح يتوفر عليها الحاسوب والطاقة الاستيعابية الكبيرة التي تؤهله للاستجابة لأي تسـاؤل، فـإن مـا هـو أساسي للتلاؤم معه هو تخزين المادة اللغوية الوثائقية والنصوص الأدبية والعلمية حتى يتمكن أي بـاحث معجمـاتي مـن أن يجـد ضـالته وفـق اسـتراتيجية مشـاريعه المعجماتيـة لإنجـاز معاجم في مستوى المرحلة، وبمـوازاة مع التطورات التكنولوجية في مجال الصناعة المعجماتية الحديثة من جهة، والأنظمة الآلية التحليلية وأجهزتها المعـلوماتية مـن جهـة أخرى على قاعدة ما حاولنا إيضاحه فيما يخص أساسيات المعطيات التي تؤدي إلى تحديد نوعي لماهية بنك النصوص (La Banque textuelle) هذا بالإضافة إلى ما سنورده عن أساسيات المعطيات النصية وأساسيات المعطيـات المعجماتيـة وأساسـيات المعطيـات القاموسـاتية، والهـدف مـن هـذا الـتحديد هـو تمييز الأجـهزة المعلوماتية أو الوثائقية المختلفة، وضبط وظائفها في أفق الإنجاز المعجماتي على اختلاف مشاريعه.

## II- أساسيات المعطيات النصية

يعـد الـنظام الآلـي للتحـليل الصرفي الذي ساهمت في إنجازه الشركة العالمية للإلكترونيات بالقاهرة قاعدة أساسية للتعامل مع مختلف النصوص. من هنا تحتل أساسيات المعطيات النصية مكانة متميزة، إذ تقدم ما يتم تخزينه ومعلمته( Informatisé ) ويكـون جاهـزاً للإجابـة عـن كـل الـتحاليل المتعـلقة بعناصـر المـلفوظات شـكلا ومضـمونا (المكتوبة أو المسـجلة) إذ أن مجمـل النصـوص التي يراد إخضاعها للتحاليل المعلوماتية تتطلب إعداداً وتهيئة وفق برامج مهيأة لهذا الغـرض. وإذا كـان لبـنوك النصوص توجه ينحصر في تدبير النسخ بحياد تام، فإن أساسيات المعطيات النصية، حسب مـا أوضـح ذلـك كـيمادا، تسـتفيد مـن مصـادر موسومة بالتوجهات العلمية للباحثين الذين خططوا لها وأنجزوها في أفق أبحاثهم الخاصة.(7)

إن كـل نـص بمجـرد مـا يصـبح قـابلا للاستخدام آليا بواسطة ما تم تهيئته من برامج وفق الغايات المحددة له، وحسـب معـايير معيـنة، يشـكل أسـاس المعطيـات النصية، وفي هذه الحالة عندما يتوفر الباحث المعجماتي على نص أو نصـوص ممعـلمة (Informatisés) بواسـطة بـنك النصـوص فـإن عـليه أن يتولى إيجاد الافتراضات النظرية ليتمكن من إنجـاز مـا يصـبو إليـه حسـب القواعـد والـتجيهزات الـتي يمـلكها، ووفـق الأهـداف المـتوخاة الـتي يريد الوصول إليها وتحقيقهـا. وأعـتقد أن بـرامج الشـركة العالميـة للإلكترونيـات قـد تمكـنت من تحديد العديد من هذه الأهداف، وما هو مطـلوب مـنها الآن هـو تقويـة أساسـيات المعطيـات النصـية مـن أجـل تكويـن قـاعدة عامـة تحتوي على مدونة مجهزة

بالإشارات ومتوفرة على الرموز التحليلية ومتضمنة برامج عيارية قابلة للتطبيق والمعالجة.

ينبغي الإشارة هنا إلى أن أساسيات المعطيات النصية لا تخص فقط البحث المعجماتي لأن طبيعتها تجعلها منفتحة على كل التخصصات العلمية وقابلة للاستثمار في شتى المجالات، وبإمكان أي باحث في مجال تخصصه استغلال، بأقصى ما يمكن من الفعالية، وثائق ضخمة ومعقدة تتضمن وقائع ومعلومات، إذ أن هذه الوثائق قد تم تسجيلها حسب قواعد نظامية، فهي بالضرورة تستجيب لكل تساؤل وحسب طبيعة التخصص. لذا فإن كل تخصص مدعو ليزود أساسيات المعطيات النصية بكل الوثائق النصية لتكون ذات طبيعة شمولية وفي مستوى المهام المطلوب إنجازها. وطبيعي أن هناك علاقة وطيدة بين بنك النصوص وأساسيات المعطيات النصية، وعلى سبيل المثال فإن الباحث المعجماتي يضطر إلى اللجوء إلى بنك النصوص لإغناء خطة عمله، وتفادي أي نقص ممكن حسب توجهات ما يريد القيام به، سواء تعلق الأمر بمعجم وسيط، أو معجم شامل، أو معجم تاريخي.

وإذا ما تم ربط وحبك العلاقة بين بنك النصوص وأساسيات المعطيات النصية، فإن ذلك من شأنه أن يوفر لنا إمكانات ضخمة، لإنجاز أهم المشاريع المعجماتية، وبأسرع ما يمكن، وجعل القارئ العربي يتوفر على معاجم مختلفة ومتخصصة هو بأمس الحاجة إليها، ويمكن الإشارة هنا إلى بعضها:

- معجم عربي شامل.
- معجم عربي حديث.
- معجم المترادفات.
- معجم الأضداد.
- معجم حياة الحيوان، وهذا المعجم يمكن أن تتفرع منه معاجم كمعجم الطيور ومعجم الحشرات ومعجم الضواري ومعجم الزواحف ومعجم الأسماك الخ...
- معجم النباتات بكل تفرعاتها.
- معاجم المصطلحات في كل التخصصات.

هذا بالإضافة إلى أن بنك النصوص، وأساسيات المعطيات النصية ، إذا ما تم توسيع نطاق التخزين، يسمحان بإنجاز مشاريع هائلة تتعلق بالتاريخ العام والتاريخ القومي، وتاريخ الفن العام والخاص بكل فروعه.

## III- أساسيات المعطيات المعجماتية

يعرف برنارد كيمادا أساسيات المعطيات المعجماتية بأنها من الأسس النصية المتخصصة وهي مهيأة ومجهزة لتحاليل نمط المعجمية، وتكمن مهمتها في الإجابة عن كل التساؤلات المتعلقة بالمفردات داخل سياقاتها. وكما يشير كيمادا ، فإنه يمكن أن نلاحظ أن التجهيزات من نمط "النصية" ضمنت بالأساس إلى حد الآن أداءات من نمط (المفرداتية) مما يعني أن الكلمات واستعمالاتها وسياقاتها هي التي تجسد أساسيات المعطيات النصية الخاضعة لمعالجة حاسوبية وفق منهجية دقيقة ومكثفة وذات مردودية عالية.(8)

إن أهم إشكال يمكن أن يواجه المعجماتي أمام هذه التقنية العالية والدقيقة، فيما يقدمه من مواد ومعلومات، هو أنه عندما يطلب مثلا كلمة « بيت » أو "مدينة" أو "حديقة" أو "أخرج" يجد نفسه أمام مئات من الأمثلة والاستعمالات والسياقات، إما منفردة، وإما مكررة، مما يفرض عليه دقة الاختيار. ويرتبط الاختيار هنا بطبيعة ما يريد إنجازه من معاجم وحسب مستوياتها وأهدافها، وهذه من المهام الملقاة على عاتقه وحده ولا يمكن للحاسوب أن يقوم بها.

هناك إذا التقنية الآلية التي تترك للمعجماتي مجالا للاختيار الخاضع للذوق والتوجهات والتحليل المعجمي، مع العلم أن إشكالية الاختيار يمكن التغلب عليها،وكما يقول كيمادا إن ذلك رهين بإيجاد المعجم الآلي ( Le dictionnaire de Machine) الذي لا غنى عنه مستقبلا لكي يوضح ويبين العناصر المفرداتية، ويعرف الرتب النحوية، ووظائف الجمل النحوية، وحيث يتم اختيار الاستشهادات مسبقا، وهناك اجتهادات متواصلة في أوروبا لإيجاد حلول عملية لهذا النوع من الإشكالات.

## IV- أساسيات المعطيات القاموساتية

يتحدد مضمون أساسيات المعطيات القاموساتية حسب برنارد كيمادا في طبيعة المعطيات المقترحة للتساؤل (عناصر مواد المعاجم) وفي دقتها حيث تمد مؤلفي الفهارس والباحثين والمعجماتيين بمنفذ سريع ومتعدد التوجهات للوصول إلى مختلف المضامين القاموساتية، إذ أن كل معجم له خصوصيته وتنوعه وسياقه الخاص الخاضع لبنية تنظيم أجزائه. ولكي تؤدي أساسيات المعطيات القاموساتية وظائفها عند المساءلة يجب أن تخضع للتصنيف وتقوية معالجتها وهي بهذا التوجه تتطور، كما أورد ذلك كيمادا،(9) كالتالي:

يتضمن تسجيل وتخزين المعاجم الموجودة (المنشورة أو التي تم تحريرها) على الحاسوب وبفضل شفرات تحليلية خاصة تكمل عملية النقل، لأن النص المنقول يعالج بطريقة المدونة الوثائقية، مما يعني أن جمع نظام الإشارة المنطقي (Système d'indication logique ) مع لغة التساؤل الملائمة تجعل بالإمكان الوصول إلى مختلف المضامين والتعرف عليها والوقوف على التباين أو التشابه ما بين معجمين والمقارنة بينهما، مما يسمح بالإضافة أو الحذف. والمعاجم التي يتم إدخالها وتخزينها حسب هذا النظام تسمح بتحاليل متعددة من بينها تحليل ميكنزمات أساليب

المعالجة النهائية. وإذا ما تطور هذا الأسلوب واستخدم في مجال اللغة العربية يمكن أن يقدم خدمات وفوائد لا حصر لها، منها أنه يسمح بمراجعة انتقائية للنسخة الأولى بواسطة أسلوب المقارنة لكل المداخل لتكملتها أو لحذف ما هو متنافر وشاذ ومتجاوز أو غير ملائم شكلا ومضمونا، وتجاوز الثغرات والهفوات، ولكي تصير متجانسة ومعيارية حسب التوجه العام لخطة المعجم المرغوب في إنجازه. ومن دون أدنى شك أن هذه الإجراءات التطبيقية تسهل عملية التصحيح والتنقيح بواسطة التحليل الآلي الصرفي ومصادر المعالجة للنص المعلم( Infomatisé) وهذا يساعد على:

- تسهيل وضع معاجم في نسخ تامة ومضبوطة.
- تهيئة طبعات جديدة.
- إمكان إعادة الطبع مباشرة دون اللجوء إلى إعادة التصحيح بالطريقة التقليدية.
- الاستفادة من المصاريف المكلفة التي كان يتطلبها إنجاز معجم جاد.
- إمكان طبع معاجم مختصرة أو خاصة بمرحلة تعليمية معينة، أو معاجم متخصصة، أو معاجم مزدوجة اللغة.

ولجعل هذه الأعمال متوفرة فمن الضروري تخزين كل المعاجم العربية قديما وحديثا وإخضاعها لهذا الغرض، أي إيجاد كتلة هامة من المعاجم على اختلاف أنواعها سواء على مستوى تعددها واتساعها، أو على مستوى طبيعتها ومضامينها لتصبح مادة معجم ما أو معاجم متنوعة، ولأنها بهذه الصورة تشكل مجموع المعطيات والمعلومات التي تقف عليها لتستغل استغلالا مزدوجا أو متعدد الأبعاد وحسب وظيفة الحاجات الضرورية لكل نوع من الإنجازات المعجماتية سواء تعلق الأمر بمعاجم أحادية أو ثنائية اللغة، وكيفما كان مضمونها لغويا أو موسوعيا،أو متخصصا. وما يرسخ هذا التوجه من الإنجاز المعجماتي يعود إلى ما يتوفر عليه الحاسوب من طاقة استيعابية وبرامج وأنظمة آلية ذات أبعاد متعددة الجوانب على مستوى الجذور والمشتقات والتعابير، أو على مستوى القضايا الصرفية والنحوية، أو على مستوى الترادف والأضداد، أو على مستوى الشواهد.

وبذلك فإن المخزون في هذه الحالة بإمكانه أن يمد الباحث في مجاله بما يحتاجه، وتلبية رغباته بالسرعة المطلوبة.

في إطار هذه المنظومة المتكاملة المعطيات، التي تعتبر أساس الاشتغال والتعامل مع اللغة، يتم التحكم في ضبط الكلمات والتراكيب وفي ضوء أنظمة قائمة على قواعد منهجية مما يسمح بالإنجاز السريع من جهة والدقة التامة، لا فيما يخص الداخل المعجمية والاشتقاقات فقط، بل وفي مجال تنوع التراكيب مما يعطي لأي معجم ينجز في ظل هذه

الشروط قيمة علمية دقيقة.

## الهــوامـش:

أنظر :

1. René Moreau et Isabell warmesson: ordinateur et lexicographie ( lexique) 2/le dictionnaire, 1983, P.U.L.p. 121.
2. B. Quemeda: Bases de données informatisées et dictionnaires (lexique) 2/le dictionnaire 1983, P.U.L. p 101-120.

نشير هنا أيضاً إلى مجموعة من الأبحاث قدمت في إطار ندوات علمية منها:

- عبد الفتاح أبو السيدة، الحاسب الآلي والترجمة، اللسان العربي، ع28، ص1987 ¡ 95.
- السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات، ندوة عقدت في الفترة من 8 إلى 12 ذي القعدة 1412هـ الموافق 10-14 مايو 1992، مكتبة الملك عبد العزيز العامة بالرياض1993/1414م.
- المؤتمر الدولي الثاني، اللغة العربية والتقنيات المعلوماتية المتقدمة، أعمال المناظرة ، الدار البيضاء 9 ¡ 8 دجنبر 1993، مؤسسة الملك عبد العزيز آل سعود للدراسات الإسلامية والعلوم الإنسانية الدار البيضاء.

3. B. Quemada p. 106.

4. نفس المرجع السابق.

5. من وثائق الشركة العالمية للإلكترونيات، " القارئ الضوئي العربي خطوة على طريق تقنية معلومات عربية " ص.3.

6. فيما يتعلق ببنوك المصطلحات والمعطيات وما كتب عنها من الناحية التقنية في بداية اشتغالها انظر: د. علي القاسمي نحو تطوير بنوك المصطلحات، اللسان العربي، ع28، ص77.

- د. ليلى المسعودي، بنوك المعطيات، اللسان العربي ع1987 ¡ 28.

قاعدة المعطيات المعجمية: المعربي، اللسان العربي، ع25، ص1985 ¡ 91.

- عامر إبراهيم قنديلجي، بنوك وشبكات المعلومات الآلية، مكوناتها، ومستلزماتها نماذج عربية وأجنبية، اللسان العربي ، ع25، ص1985 ¡ 109.

- محمد محمد أمان بنوك المعلومات، تونس، المنظمة العربية للتربية والثقافة والعلوم 1983.
- وفيما يتعلق ببنك الكلمات انظر مجلة :

La banque des mots, revue de terminologie française publiée par le conseil international de la langue française.

7. B. Quemada, p 110.111.

8. نفس المرجع السابق ص111-112.

9. نفس المرجع السابق، ص115.